



Optimizing LONTALK Response Time

August 1991

LONWORKS™ Engineering Bulletin

Introduction

Several tools are available to the LONWORKS network designer to optimize response time. The LONTALK™ protocol supports a variety of optional messaging services and collision detect hardware. The features of these services and their effect on LON® response time are discussed.

Media Access Delay & Offered Traffic

When a node on a network tries to send a message, it must first wait for the medium to be idle¹. The time delay between when a node queues the packet to send and the time that it is actually sent on the network plays a role in the response time of that message. This delay is known as the media access delay. As the offered traffic (the total number of packets per second offered for transmission by all the nodes on a channel) on a given channel of a network increases, the media access delay increases. In networks loaded to near-capacity, when many nodes are trying to send messages, this delay can be significant. In designing a network to meet specified response times, the worst-case offered traffic must be considered and designed for.

In order to assess the impact of media access delay, we need to first consider the amount of time a typical message spends on the media. For example, consider a network design where the average packet size is 112 bits (14 bytes) and the average number of bits between each packet is 48. Also, suppose that the channel runs at 78 Kbps. Each packet cycle takes 160 ($= 112 + 48$) bit times on the channel. Thus, the maximum rate at which packets can be sent on this channel is 487 packets ($= 78000/160$) per second, and each packet cycle requires 2 ms.

Now consider a node on this network that has a maximum response time requirement of 50 ms. In order to design this network to meet this response time requirement, we must set a bound on the offered traffic to limit the media access delays. At 78 Kbps with 160 bit packet cycles it takes this network 50 milliseconds to send 25 packets. If we further assume that the packets on the channel may need to be sent at any time, including all at once, then, in order to always meet the requirement of 50 milliseconds without using any special priority features, the channel should never have an offered traffic greater than 25 packets per second. As this example illustrates, the response time requirements of nodes on a network

¹ Details are described in the LONWORKS Engineering Bulletin entitled *Enhanced Media Access Control with Echelon's LONTALK Protocol*.

must be used to design the network and ensure that the offered traffic does not cause unacceptable media access delays

Optional Priority

The LONTALK protocol supports an optional priority messaging service that enables intelligent nodes to obtain optimal response times. When a node tries to access the communication medium, priority messages are given earlier access than non-priority messages. In the LONTALK protocol, the effectiveness of priority depends on all nodes preparing to send a message being synchronized on the end of the previous packet. When the sending nodes are synchronized, then the priority time slots remain uncontested. Priority assignment has two limitations: (i) only a limited number of nodes may be assigned priority; and (ii) reserving time slots for nodes of different priorities pre-allocates bandwidth. For Echelon's 78 Kbps twisted-pair transceiver, each priority slot is 8 bit times wide. Using the 1.25 Mbps twisted-pair transceiver the priority slots are 21 bit times wide. In effect, each packet length is increased by this amount thus using up bandwidth.

Collision Detection

A collision is defined as the event when two or more nodes access the communication medium at almost the same time, resulting in mutual interference among their electrical signals on the transmission medium. Packets involved in a collision cannot be successfully received, which results in a delay in response time.

The media access algorithm used by the LONTALK protocol minimizes the number of potential collisions by randomizing access to the medium. Details are described in the Echelon Engineering Bulletin entitled *Enhanced Media Access Control with Echelon's LONTALK Protocol*. The LONTALK protocol also encodes an estimate of the channel backlog in each packet so that the number of randomizing slots can be dynamically adjusted depending on network traffic. Randomizing slots reduces the probability of collisions, however it does not completely eliminate them. Further, as network load increases, the probability of collisions increases.

Collision detection hardware at a transmitting node informs the transmitting node shortly after a collision occurs about the need to retry. The NEURON® CHIP firmware is configured to detect a collision at the end of the packet preamble and the end of a packet transmission. In the absence of such hardware, the transmitting node only learns about the unsuccessful transmission from a lack of an explicit or implicit positive acknowledgement from the receiving node, which may take much longer. Thus, collision detection is very useful for nodes requiring fast response time.

Implementing collision detection requires media-specific hardware solutions. Echelon has designed the hardware to implement collision detection on 78 Kbps and 1.25 Mbps twisted-pair transceivers.

Network and Application Input Buffers

In a LON system, every node connected to a channel receives every packet transmitted on the channel and checks whether the packet is addressed to it. Thus the speed at which a NEURON CHIP discards packets not addressed to it determines the upper limit for the number of packets per second on a channel. If packets arrive at a node faster than the node can process the packets and free up the input buffers, packets will be lost. This has no effect if the packet was not addressed to the node in the first place. However, if the lost packet was addressed to the node, response time is affected.

Once a packet has been received in a network input buffer, its destination address is examined to see if the packet is addressed to the node. If it is, then the packet is copied into an application input buffer, and the network input buffer is freed. If no application input buffer is available (because the application is not processing its packets fast enough), then the packet is discarded and no acknowledgement is sent.

Output Buffers

The number of output buffers becomes a factor in response time when a node can generate outgoing packets faster than the network can take them. What happens in this case, is that several outgoing packets back up in the output queue. The NEURON CHIP is quite capable of generating over 300 packets per second. At lower data rates or during network saturation, this level of packet generation can easily overwhelm the ability of the network to accept packets. For this reason, it is always important to have an idea of the maximum number of messages your application may generate per second as this dictates the network data rate required.

Messaging Services

The LONTALK protocol offers four basic types of message service:

- acknowledged
- request/response
- unacknowledged repeated
- unacknowledged

These are described in the LONTALK Protocol Application Note. There are a number of tradeoffs relating to network efficiency and response time in selecting a message service. The effect of the message service timers on response time is discussed.

Transaction Timer

When the acknowledged message service is used, the transaction timer in the NEURON CHIP controls the time allowed for an acknowledgement to arrive. By adjusting the transaction timer, the time delay before messages are retried (due to lack of acknowledgement) can be optimized, hence improving response time. A retry is initiated when the transaction timer on a node expires without an acknowledgement being received. The retry count controls the number of times that a node will retry to send a message.

The network processor in the NEURON CHIP is responsible for checking for the acknowledgement. To keep the number of retries to a minimum, this timer must be set high enough to accommodate the round trip delay of sending a message and getting an acknowledgement back. On a network running at 78 Kbps or 1.25 Mbps, where the source and destination(s) are all on the same channel, the transaction timer can be set to a low value such as 64 ms or 98 ms. In the case where the packet has multiple destinations, the longest path must be used to calculate the timer value. A node will not try to initiate retransmission until its transaction timer expires.

Repeat Timer

The unacknowledged repeated message service is used when a large number of nodes must be addressed and you want to ensure that the message gets through. With a large group, an acknowledged message would cause too much traffic. The repeat timer can generally be set much lower than the transaction timer. It specifies the interval between repeats of the initial packet. The initial packet is repeated as many times as is specified by the repeat count, from 1 to 15 times. This timer should be long enough for the receiving node(s) to overcome any short term buffer shortages.

Receive Transaction Timer

This timer is used for detecting duplicate messages. Upon receipt of a message addressed to a LONWORKS node, the protocol firmware attempts to allocate a receive transaction buffer if the packet represents a new transaction. Failure to allocate a new transaction record when it is needed causes the incoming packet to be discarded, and not to be acknowledged.

A new transaction is defined as one for which there is no transaction record containing the source address of the packet and a matching transaction ID. When this timer expires, the transaction record is deleted, and any new packet from the same source address will cause a new transaction to be created. If this timer is set for too short a time, such an additional packet (a retry) with the same source address and the same transaction ID would not be detected as a duplicate, but rather it would be acted upon again. On the other hand, if this timer is set for too long a time, transaction records tend to be allocated for a long time which could exhaust the

supply of available records. This, in turn, would cause the node to discard packets because it cannot allocate a receive transaction record.

Application Response Time

The application response time indicates the time it takes an application in one node to see a change in status, and form a packet to send, plus the time for the application in the receiving node to process the packet and perform the control action.

The application in a LONWORKS node is programmed in NEURON C. The application scheduler is event-driven. The application response time is influenced by the the number of when clauses, the complexity of the conditional expressions within the when statement, and the execution times of the when clauses. In addition, the number of software timers declared, the number of address table entries, the number of network variables and message tags all affect response time. For most programs, the application response time can be ignored. If, however, your application program is near the maximum for any of these items, some measurements should be done.

Conclusion

The LONTALK protocol provides a variety of tools and optional services to optimize LON response time. The response time requirements of nodes on the network must be considered when selecting the type of message service to use: priority, acknowledged, unacknowledged/repeated, and also when deciding whether collision detection hardware should be included in the node design.

Disclaimer

Echelon Corporation assumes no responsibility for any errors contained herein.
No part of this document may be reproduced, translated, or transmitted in any form without permission from Echelon.

© 1991 Echelon Corporation. ECHELON, LON, and NEURON are U.S. registered trademarks of Echelon Corporation. LONMANAGER, LONBUILDER, LONTALK, LONWORKS, 3150, and 3120 are trademarks of Echelon Corporation. Patented products. Other names may be trademarks of their respective companies. Some of the LONWORKS TOOLS are subject to certain Terms and Conditions. For a complete explanation of these Terms and Conditions, please call 1-800-258-4LON.

Echelon Corporation
4015 Miranda Avenue
Palo Alto, CA 94304
Telephone (415) 855-7400
Fax (415) 856-6153

Echelon Europe Ltd
105 Heath Street
London NW3 6SS
England
Telephone (071) 431-1600
Fax (071) 794-0532
International Telephone + 44 71 431-1600
International Fax + 44 71 794-0532

Echelon Japan K.K.
AIOS Gotanda Building #808
10-7, Higashi-Gotanda 1-chome,
Shinagawa-ku, Tokyo 141, Japan
Telephone (03) 3440-8638
Fax (03) 3440-8639

Part Number 005-0011-01